

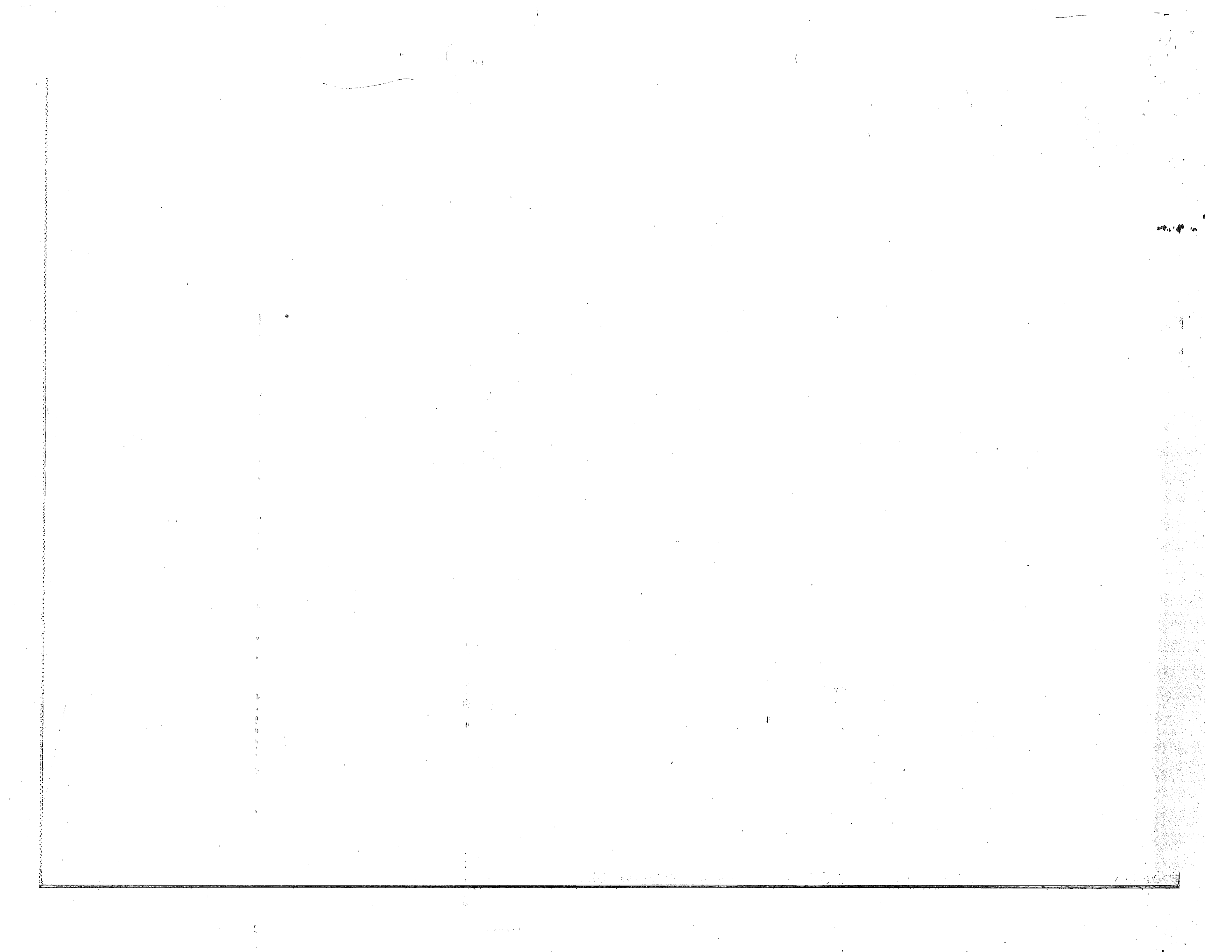
Baranger (A)

Denis Baranger

CONSTRUCCION Y ANALISIS DE DATOS

Una introducción al uso de técnicas cuantitativas en la Investigación Social

Editorial Universitaria



CAPITULO I

LA ESTRUCTURA DE LOS DATOS

«Otra peculiaridad de los datos científicos es que, pese a la etimología de la palabra 'dato', no son nada dado, sino que hay que producirlos, y muchas veces laboriosamente». (Bunge, 1979)

«¿ Qué es un dato?», es la pregunta a la que quisiéramos responder en este primer capítulo. Intentaremos mostrar así como la estructura de cualquier dato es esencialmente el producto de un acto clasificatorio, acto cuya simplicidad es sólo aparente. Lejos de sernos inmediatamente "dado" —como lo sugiere engañosamente la etimología—, el dato presupone toda una serie de operaciones que se llevan a cabo en forma simultánea dentro de un sistema conformado por múltiples dimensiones. De este modo, el dato más elemental ya conlleva toda una serie de elecciones teórico—metodológicas. Y es que desde el punto de vista que se sostiene aquí, el dato es eminentemente racional: no se "ve" más que lo que ya se está dispuesto a ver.[1]

Comenzaremos procediendo a un ejercicio de disección de la naturaleza del dato, mediante el cual estaremos en condiciones de aislar sus elementos constitutivos. La exposición algo dogmática de una serie de definiciones nos permitirá establecer las relaciones sistemáticas por las que se vinculan unos a otros los conceptos de 'sistema' y 'unidad de análisis', 'propiedades' y 'medición', 'variables' y 'valores'. Para ello, nos valdremos de ejemplos muy sencillos y deliberadamente no sociológicos. Al término de este trabajo analítico, el dato deberá aparecer como el producto de la imbricación de dos estructuras: la que permite atribuir una propiedad a un objeto, y la que posibilita estructurar sistemáticamente conjuntos de propiedades.

---

1 Desde la tradición epistemológica francesa (Cf. Bachelard, 1972 y 1986) esto es una obviedad. En la vereda opuesta, el positivismo lógico (Cf. Ayer, 1965) sostenía la posibilidad de un lenguaje observacional puro dentro del cual era factible formular enunciados que permitieran dirimir acerca de la verdad—falsedad de una hipótesis. En el mundo anglosajón, hubo que esperar hasta Kuhn (1969) para que se pusiera radicalmente en cuestión la reconstrucción positivista.

1. UNIDADES DE ANÁLISIS, VARIABLES, VALORES

Según nos dice Galtung,[2] todo dato hace referencia a una estructura constituida por tres elementos: unidad de análisis, variable y valor. Cualquier dato aislado consistirá pues en...

a) una **unidad** de análisis que

b) en una **variable** específica presentará

c) un determinado **valor**.

Lo que constituye el dato son estos tres elementos considerados conjuntamente con las relaciones que mantienen entre sí. Sea cual sea nuestro objeto de estudio, si queremos sostener cualquier proposición empírica acerca de él, se lo deberá concebir en términos de esa estructura tripartita. 'Unidad', 'variable' y 'valor' son todos términos que denotan determinados tipos de conceptos.

Una sencilla aproximación desde la lógica puede permitirnos echar algo más de luz sobre su significado, partiendo de las categorías clásicas de 'sujeto' y 'predicado'. La idea más elemental de lo que es un dato puede ejemplificarse a partir de un enunciado muy simple: «ese sillón es rojo». Se puede reconocer en este enunciado un sujeto ("sillón") y un predicado ("rojo"); también podemos decir que al hacer esta afirmación estamos denotando una determinada **propiedad** de un objeto. Lo mismo sucede cuando sostengo, por ejemplo, que "Rómulo es argentino": 'argentino' es una propiedad que se predica del sujeto de la oración ("Rómulo"). Ahora bien, las fórmulas «ese sillón es rojo» y «Rómulo es argentino» consisten simplemente en enunciados que se refieren respectivamente a diferentes proposiciones.[3] Por lo tanto, es posible establecer desde ya que:

un dato se expresa en una proposición[4]

Torgerson (1965) propuso el uso del término 'sistema' para denotar todo tipo de objetos o de "cosas"; así, en nuestros ejemplos, tanto 'ese sillón' como 'Rómulo'

---

[2] Cf. Galtung, 1968: I, cap. 1.

[3] Es importante tener en cuenta esta distinción que establece la Lógica. Así, «ese sillón es rojo» y «*ce fauteuil est rouge*» son ciertamente enunciados diferentes, pero ambos hacen referencia a la misma proposición. Según Copi, «Se acostumbra usar la palabra 'proposición' para designar el significado de una oración declarativa (...) Una oración declarativa forma siempre parte de un lenguaje determinado, el lenguaje en el cual es enunciada, mientras que las proposiciones no son propias de ninguno de los lenguajes en los cuales pueden ser formuladas» (1962: 21).

[4] Según Bunge, «Las ideas que expresan el resultado de una fase de observaciones son un conjunto de datos. Un dato es una proposición singular o existencial como, por ejemplo, 'Se inyectó a la rata # 13 1 mg de nicotina el primer día'» (1979: 742)

serían sistemas. Es evidente que, por definición, la variedad de sistemas que pueden distinguirse no tiene límites: una célula es un 'sistema', y también lo son un auto, un alumno, un libro o una galaxia. A su vez, cada sistema puede ser caracterizado por toda una serie de propiedades:



un **auto** en particular podrá ser *rojo,*
caro,
veloz,
antiguo,
etc.

Sin embargo, observemos desde ya que no todas las propiedades son aplicables a todos los sistemas. No tendría sentido pensar en un libro 'veloz', así como carecería de significado hablar de una galaxia 'cara' (al menos para nosotros, aquí y ahora). Esta es una simple consecuencia del hecho de que un dato deba expresarse en una proposición, vale decir en algo que necesariamente habrá de ser verdadero o falso. Así el enunciado «este libro es veloz» no haría referencia a ninguna proposición.[5]

Cuando a partir de la observación de un sistema predicamos una propiedad, podemos decir que estamos realizando una **medición**.[6] Lo fundamental es que los sistemas no son jamás medidos en sí mismos sino que lo que medimos son siempre propiedades de los sistemas; es así que estamos en condiciones de producir una nueva definición:

un dato es el producto de un procedimiento de medición, y medir supone predicar una propiedad.

Encontramos ejemplos de datos muy simples en enunciados del tipo:

«Este **auto** es................*rojo*.»,
«Esa **galaxia** es...............*lejana* »,
«Este **libro** pesa...........*358 gramos* »,
etc.

Es desde este punto de vista que se puede afirmar que el conocimiento científico consiste en la identificación de sistemas de determinadas clases, en la medición de sus propiedades, y en el establecimiento de relaciones entre dichas

---

[5] «Para los fines de la lógica, puede definirse una proposición como algo de lo cual es posible afirmar que es verdadero o falso» (Cf. Cohen y Nagel, 1968: I, 41). Convendrá notar que esto abre la posibilidad de que existan **datos falsos**: así contrariamente a lo afirmado con anterioridad, lo cierto es que «(mi amigo) Rómulo es uruguayo.»

[6] Al menos implícitamente; puesto que la simple afirmación de la propiedad es lógicamente equivalente a la negación de la no—propiedad.

propiedades. Por lo demás, es obvio que en ciencias sociales no nos interesan todos los sistemas, ni tampoco todas sus propiedades. De hecho, sucede que precisamente puede entenderse que lo que diferencia a las disciplinas científicas unas de otras son en gran medida los tipos de sistemas de los que se ocupan y las propiedades de dichos sistemas que toman en cuenta.[7]

Por otra parte, basándonos en determinadas propiedades podemos referirnos a **clases** de sistemas:

las **galaxias lejanas**,
los **autos 0 kilómetro**,
los **libros de Sociología**,
los **alumnos de Metodología**,
etc.

También de estas clases de sistemas podemos predicar propiedades:

«Los **autos 0 km.** son *caros* »,
«Las **galaxias lejanas** son *difíciles de observar*»,
«Los **libros de Sociología** son *aburridos* »,
«Los **alumnos de Metodología** son *sabios* »,
etc.

Así como existen autos de diversos colores, también existen lápices, flores, libros y cielos de distintos colores. Todos concordaremos sin embargo en que un auto no es ni un lápiz ni una flor, ni un libro ni un cielo. Si queremos saber qué es un auto podemos recurrir al Larousse usual: « Vehículo que camina movido por un motor de explosión». Todos coincidiremos en que un lápiz no "camina", en que una flor no es un "vehículo", y en que los libros y cielos no son "movidos por un motor de explosión". Es imposible que confundamos un auto con un lápiz, porque de algún modo poseemos **definiciones** de estos objetos; sabemos que un auto se define por determinadas características: tener cuatro ruedas, un volante, asientos y una carrocería, y ser capaz de llevarnos con relativa —y cada vez menor— seguridad de un punto a otro de la superficie terrestre, rodando por los caminos. En otras palabras, el sistema 'auto' se caracteriza por presentar determinadas propiedades, a saber:

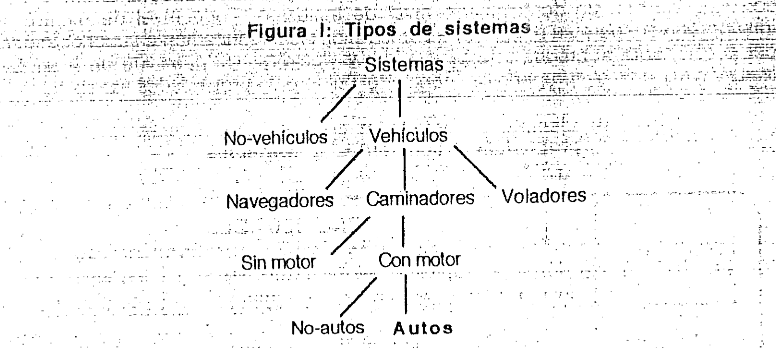
"ser *vehículo* ",
"*caminar* ",
"ser movido por un *motor a explosión* ",
"tener *cuatro ruedas* ",
etc.

7 O, para decirlo de otra manera, las ciencias se diferencian por sus "**objetos**".

Se pueden imaginar fácilmente sistemas que no satisfagan algunas de estas propiedades: una avioneta Piper o una lancha no "caminan", una bicicleta no es "movida por un motor de explosión", etc.

Lo que queremos remarcar, entonces, es que no solamente existen sistemas de los cuales es posible predicar propiedades, sino que los sistemas mismos son susceptibles de ser definidos en última instancia a partir de una determinada combinación de propiedades. Mucho "antes" de ser *rojo*, un 'auto' es **auto** en tanto presenta determinadas propiedades. Podríamos pensar que todos los sistemas se dividen en dos **clases**: "Vehículos" y "No—vehículos"; que algunos de los vehículos "caminan" mientras otros no lo hacen, y que sólo algunos de estos últimos son "movidos por un motor de explosión":

Figura I: Tipos de sistemas



A partir de las nociones de "sistema" y de "propiedad", podemos establecer los conceptos de "universo" (y, correlativamente, el de "muestra"),[8] "unidad de análisis", "variable", y "constante". Estos términos cobrarán un sentido específico cuando se los considere dentro de una estructura definida en un determinado nivel de inclusión. Así, podemos reconocer en el "**Universo**" el nivel de mayor inclusión posible, que se compone de todos los sistemas existentes, de todos los que han existido y existirán; en este nivel se incluyen también —al menos potencialmente— todas las propiedades posibles. Ahora bien, apelando a determinadas propiedades podemos definir **sub**-universos —o, simplemente, "**universos**"— dentro de este "Universo" mayor en un número ilimitado. Por ejemplo, es posible definir el universo de los "vehículos". Los vehículos son sistemas que permiten desplazar a otros sistemas (y también a sí mismos) de un punto a otro del espacio: se los define por poseer la propiedad de ser "medios de locomoción".[9] Se los puede distinguir así de todos los otros sistemas en los que no reconocemos la propiedad de funcionar como medios de locomoción. Todos

8 Cf. *Infra*, capítulo III.
9 Larousse *dixit*.

los vehículos comparten esta propiedad de ser "medios de locomoción": para esta clase de sistemas, el ser "medio de locomoción" es una **constante**. Notemos que en el nivel de inclusión inmediatamente superior —el Universo— no todos los sistemas son "medios de locomoción": en ese nivel superior dicha propiedad funciona como una **variable**.

Podemos continuar definiendo universos de menor nivel de inclusión. Así tendremos el universo de los vehículos "caminadores" (que se diferencia de los universos de los vehículos "voladores" y "navegadores"); a su vez, dentro del universo de los vehículos caminadores, definimos el de los "movidos a motor", y dentro de éste, el universo de los autos. La Figura II. permite apreciar cómo cada universo [10] está definido por una serie de propiedades constantes y se corresponde con un determinado nivel de inclusión. Todo nivel de inclusión inferior requiere para su definición a partir del nivel inmediatamente superior tomar como constante al menos una propiedad; y viceversa: al pasar de un nivel de inclusión menor a otro superior aumentará el número de propiedades variables.

Figura II: Ejemplo de definición de universos en base a propiedades constantes y variables

UNIVERSOS	PROPIEDADES			
	Medio de locomoción	Medio en que se desplaza	Medio de impulsión	Número de ruedas
Sistemas	Variable	Variable	Variable	Variable
Vehículos	Constante	Variable	Variable	Variable
Vehículos caminadores	Constante	Constante	Variable	Variable
Vehículos caminadores a motor	Constante	Constante	Constante	Variable
Autos	Constante	Constante	Constante	Constante

En cualquier nivel de inclusión que se tome, las propiedades aparecerán jugando entonces un doble papel: algunas de ellas, tomadas como constantes, estarán definiendo un universo, mientras que otras funcionarán como variables y podrán ser objeto de investigación. Podemos ahora introducir la idea de que todos

10 Excepto, claro está, el primero de todos: el "Universo de todos los sistemas".

los sistemas que corresponden a un nivel de inclusión dado conforman un universo, y que cada uno de estos sistemas que componen dicho universo es susceptible de ser considerado como una unidad de análisis. Ello nos permite definir:

Una unidad de análisis es un sistema definido por presentar determinadas propiedades, algunas de ellas constantes (las que definen su pertenencia a un universo compuesto por todos los sistemas que presentan esas mismas propiedades) y otras variables (las que podrán ser materia de investigación dentro de ese universo).

Disponemos ahora de una primera definición de lo que es una 'variable', entendida como un determinado tipo de propiedad, a partir de los conceptos de 'unidad de análisis' y de 'universo'. En lo que sigue, nos proponemos continuar elucidando este concepto de 'variable', para lo cual deberemos introducir una nueva idea: la de 'valor'.

La siguiente definición de Galtung es un buen punto de partida para esclarecer las relaciones que median entre los tres elementos constitutivos del dato:

«Dado un conjunto de unidades, un valor es algo que puede predicarse de una unidad, y una variable es un conjunto de valores que forma una clasificación.»[11]

De acuerdo a lo que se ha venido exponiendo, no debe haber ninguna dificultad en reconocer de inmediato que un 'valor' es simplemente una propiedad. Ahora bien, hemos dicho que una variable también es una propiedad; ¿Significa esto que 'valor' y 'variable' son dos términos que denotan el mismo concepto, el de 'propiedad'?. En absoluto.

Cuando enunciamos «Este **auto** es *rojo* », ¿ En qué sentido es posible hablar de "medición"? Solamente en tanto esta propiedad '*rojo*' se encuentre definida por su lugar dentro de una estructura que le confiera un sentido. Así la estructura más simple en la que podemos pensar se basa en una relación simple de oposición:

rojo versus *no—rojo*

De este modo afirmar que el auto es rojo equivale a afirmar que dicho auto **no** es no—rojo: la afirmación de la propiedad es al mismo tiempo la negación de la no—propiedad. Podríamos introducir un término para denotar esta propiedad —el de '*rojidad.*'— que denotaría una variable, entendida ahora como una característica

11 Cf. Galtung, 1968: I, 78.

susceptible de adoptar diferentes valores ("sí" y "no"). Cualquier unidad de análisis de la cual se pudiera predicar con sentido la rojidad presentaría en esta variable el valor "sí" o el valor "no". Así como podríamos medir la longitud de un auto («este auto mide 427 cm.»), así también mediríamos su 'rojidad':

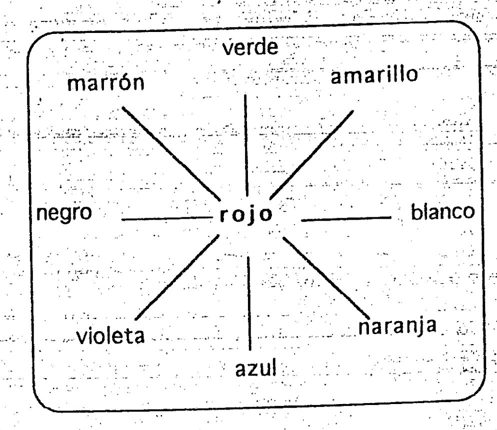
> «Este **auto** presenta el valor "sí" en la variable *rojidad*»
> (o, para decirlo más simplemente: «Este auto es rojo»).

De este modo, podemos establecer que:

una variable es un conjunto estructurado de valores[12]

Ahora bien, el término 'rojo' también puede definirse por su pertenencia a una estructura algo más compleja; dentro de esta estructura, el 'rojo' se definirá como tal por **no ser** ni marrón, ni verde, ni amarillo, ni blanco, ni naranja, ni azul, ni violeta, ni negro. 'Rojo' puede aparecer ahora como un valor de una variable a la que podríamos denominar '*color*', y en la que se opone a toda una serie de otras propiedades que constituyen otros tantos valores posibles de esa variable.

Figura III: Ubicación de "rojo" dentro de una estructura de oposiciones



Los nombres de colores incluidos en la Figura III constituyen "un conjunto de valores que forma una **clasificación**", vale decir, lo que Galtung define como

12 Para Frege, uno de los padres de la moderna Lógica, «Una variable es un signo que, en lugar de tener un designado fijo, "recorre" un espectro de posibles designados, y la determinación de un designado particular consistirá en la asignación de un valor para esa variable» (Cf. Lungarzo, 1986: I, 36).

'**variable**'. 'Variable' y 'valor' aparecen pues como conceptos bien distintos, aun cuando se encuentren estrechamente relacionados. Un valor es parte de una variable, no existe sino por las relaciones que guarda con los otros valores que componen esta variable. A su vez, una variable no es otra cosa que el conjunto de los valores que la conforman y de las relaciones que éstos mantienen entre sí.

Las variables correctamente construidas deben necesariamente satisfacer los requisitos demandados a cualquier clasificación: los valores que las compongan deberán en cada caso ser **exhaustivos** —en su conjunto—, y **mutuamente excluyentes**. Así como la hemos construido, nuestra variable 'color' es imperfecta: un auto rosado no encontraría ubicación en ningún valor; nuestra clasificación no es exhaustiva, y requeriría del agregado de otros valores (o, **categorías**). En cambio, sí resultan mutuamente excluyentes las categorías (o valores) que la componen: cualquier persona de cuatro o más años en nuestra cultura dispone de un sistema de reglas que le permite aplicar correctamente estos valores.

2. NIVELES DE MEDICION

Puesto que las variables no consisten en otra cosa que en un conjunto de valores vinculados unos a otros por relaciones de diferente tipo, una base fundamental para la clasificación de las variables es la que se encuentra en el tipo de relaciones que mantienen entre sí sus valores. Dicho de otro modo, las variables pueden ser muy diferentes en cuanto a su estructura interna. La importancia de esta distinción es crucial, porque las estructuras constitutivas de las variables determinan totalmente el tipo de operaciones a las que pueden ser sometidas.

Remitiéndonos a la clásica distinción entre ciencias **fácticas** y ciencias **formales**,[13] recordemos que Bunge destaca el rol instrumental de las ciencias formales; es en este sentido que se puede decir que todas las ciencias fácticas recurren a la Lógica y todas tienden al uso de las Matemáticas.[14] La razón principal para ello es que las matemáticas pueden ser vistas como un arsenal de modelos de distinto tipo aplicables a la materia de cada una de las disciplinas que versan sobre la naturaleza o la sociedad. A partir de la formalización de determinados aspectos que hacen a los objetos reales de las ciencias, el uso de las matemáticas habilita para operar simbólicamente sobre los objetos de conocimiento, con toda la ganancia en eficiencia y precisión que ello entraña para la práctica científica.

13 Cf. Bunge, 1968: capítulo 1.

14 Puede cuestionarse este rol "instrumental" de las matemáticas con respecto a las ciencias fácticas. Así, según Althusser (1985 : 31 y ss.) el papel de las matemáticas en la Física era **constitutivo** antes que meramente instrumental: la Física es impensable sin el cálculo. Empero, aún desde esta perspectiva, las matemáticas son distintas de la Física, en tanto existen con independencia de ésta y se utilizan en otras disciplinas.

Ahora bien, al decir de Blalock, «el empleo de un determinado modelo matemático supone que se ha alcanzado un determinado nivel de medición».[15] Es decir que la posibilidad de hacer uso de modelos matemáticos está dependiendo de que se hayan desarrollado procedimientos de medición satisfactorios. Desde este punto de vista, no es del todo errónea la idea que lo que distingue a las ciencias sociales en general de las ciencias de la naturaleza globalmente consideradas son los distintos niveles de medición en que operan predominantemente unas y otras.[16] Por supuesto, existen también diferencias —y no menores— dentro de estos dos grandes conjuntos de disciplinas. Así, la Economía se nos aparece como una ciencia social esencialmente cuantificadora, que nos habla permanentemente en **números**: horas de trabajo, toneladas de producto, valor en pesos, etc.; ha avanzado mucho más que la Sociología en dirección hacia la cuantificación.[17]

A partir del clásico trabajo de Stevens de 1951,[18] es corriente considerar que existen cuatro distintos niveles de medición: nominal, ordinal, intervalar y racional.[19] Algunos autores definen a la medición como el proceso de asignar números para representar cantidades,[20] y excluyen de este modo el nivel nominal. Otros, a cuyo criterio nos atendremos, no consideran que la cuantificación sea una condición necesaria para poder hablar de 'medición'. Así, por ejemplo, se puede definir 'medida' como «cualquier procedimiento empírico que suponga la asignación de símbolos, de los cuales los numerales son solamente un tipo, para los objetos o sucesos que concuerdan con las reglas previamente establecidas».[21] Lo cierto es que la operación de medición será posible solamente en tanto y en cuanto se pueda determinar una cierta correspondencia entre las características discernibles de nuestros objetos de estudio y las propiedades de los modelos matemáticos que les sean aplicables. Según resumen Padua y Ahman: «La medición, de hecho, corresponde a una serie de teorías conocidas como **niveles de medición** : detrás de cada uno de los distintos niveles de medición están operando una serie de principios lógicomatemáticos, que van a determinar o no el isomorfismo entre un concepto y el nivel de medición apropiado

15 Cf. H. Blalock, 1966: 29.

16 Cf., por ejemplo, Torgerson (1965).

17 Como lo demuestra *a contrario* el estudio de las sociedades primitivas, esta posibilidad de cuantificación de la Economía está directamente ligada a una característica histórica de su objeto real: la aparición del **dinero** como medio de intercambio generalizado y la reducción de todas las mercancías, fuerza de trabajo incluida, a una base común producida por la generalización del modo capitalista de producción.

18 Reproducido en la compilación de Wainerman (1976).

19 Coombs (en Festinger y Katz, 1976) distingue seis niveles o escalas.

20 Así, Blalock define: « Por medición, nos referimos al proceso general por el que se asignan números a objetos de tal modo que queda establecido qué tipo de operaciones matemáticas pueden ser utilizadas legítimamente, dada la naturaleza de las operaciones físicas que se han usado para justificar o racionalizar dicha asignación de números a objetos» (1982: 11).

21 Cf. C. Selltiz *et al.*; 1965: 215.

(...) se busca que el modelo matemático sea isomórfico con el concepto; esto es, que la 'forma' del modelo sea idéntica a la 'forma' del concepto» (Padua, 1979: 154).

Al hablar de **niveles** de medición se está aludiendo ya a la existencia de una jerarquía entre los modelos matemáticos aplicables a los objetos; y es que, en efecto, los hay más poderosos que otros. Presentaremos a continuación las características básicas de los cuatro niveles distinguidos por Stevens, comenzando por el más bajo de ellos.

Nivel nominal

La operación más básica y a la vez más sencilla de toda ciencia es la **clasificación**. Al clasificar, agrupamos nuestras unidades de análisis en clases mutuamente excluyentes basándonos en sus semejanzas y diferencias. De este modo obtenemos una escala nominal, que corresponde al nivel más bajo de medición, y que es a su vez una condición necesaria que deben satisfacer los niveles más elevados. El requerimiento fundamental para construir una escala nominal es que se puedan distinguir dos o más categorías significativas y que se establezcan los criterios según los cuales los individuos, grupos, objetos o respuestas serán incluidos en una o en otra categoría. Las categorías —exhaustivas y mutuamente excluyentes— no mantienen otra relación entre sí que la de ser distintas unas de otras: sobre todo no existe implicación alguna de que puedan estar representando una cantidad mayor o menor de la propiedad que se está midiendo.

En ciencias sociales son muchas las propiedades que se encuentran definidas nominalmente: sexo, estado civil, nacionalidad, religión, afiliación partidaria, etc. Es cierto que nada prohibe la utilización de números para identificar las diferentes categorías, codificando por ejemplo 'masculino' como '1' y 'femenino' como '2'; pero es esencial comprender que aquí los números no funcionan como tales, en el sentido de que no autorizan el empleo de las operaciones aritméticas usuales. Así, carecería de sentido sumar los códigos que presenten en la variable 'sexo' todos los integrantes de un grupo familiar, del mismo modo que resultaría disparatado sumar sus números de documento de identidad; pero sí en cambio la suma de sus ingresos mensuales en pesos tendría un significado claro. En el nivel nominal, si utilizamos números les hacemos cumplir la función de meros **nombres** que nos permiten identificar las categorías designándolas de modo diferente. La operación de medición queda por lo tanto limitada a referir una observación a una clase o categoría para luego poder contar cuántos casos caen dentro de cada categoría. Esto limita severamente las operaciones estadísticas utilizables que no van más allá del número de casos, el cómputo de porcentajes, el modo, el cálculo de algunos coeficientes de asociación en tablas de contingencia (Q, Phi, etc.) y de la chi-cuadrada.

Nivel ordinal

En este nivel aparece la idea de orden. No sólo se puede distinguir entre distintas categorías, sino que es además posible afirmar si una unidad de análisis posee en mayor, en igual o en menor grado que otra la propiedad que se está midiendo. Así, por ejemplo, en la escala de la jerarquía militar no solamente podemos decir que un capitán, un teniente y un sargento son diferentes entre sí, sino que dichas categorías pueden ordenarse con respecto a la **autoridad** que reviste cada uno de esos grados:

Capitán > Teniente > Sargento

Pero lo que no podemos determinar con exactitud es cuánta más autoridad tiene un capitán que un teniente, ni si la diferencia entre éstos es menor o mayor que la que media entre los grados de sargento y de teniente. « Una escala ordinal define la posición relativa de objetos o individuos con respecto a una característica, sin implicación alguna en cuanto a la distancia entre posiciones».[22] La escala de dureza de Mohs, que se aplica a los minerales, es un clásico ejemplo de escala ordinal en las ciencias de la naturaleza.

En ciencias sociales las variables medidas en el nivel ordinal son muy numerosas: los índices de nivel económico—social (NES), el prestigio ocupacional, el grado de radicalismo político o de nacionalismo, la religiosidad, etc. Así, es frecuente que los NES se ordenen de 1 a 7, sin que ello signifique que los números funcionen plenamente como tales, puesto que sólo representan relaciones del tipo 'mayor que' o 'menor que'. Puede afirmarse que el NES 7 es distinto del NES 6; puede decirse también que el NES 7 es más alto que el 6; pero no podemos saber si la diferencia entre ambos es **igual** a la que separa el NES 2 del 1, ni tampoco podremos sostener que el valor 4 representa el **doble** de status socioeconómico que el 2. No obstante ello, se amplían en este nivel las medidas estadísticas utilizables; además de las aplicables a escalas nominales, se agregan la mediana, los cuartiles y percentiles, coeficientes de correlación de rangos como el de Spearman, etc.

Nivel intervalar

Cuando se alcanza a construir una escala de **intervalos iguales**, no solamente es posible diferenciar y ordenar las unidades de análisis, sino que también se puede especificar la distancia que separa a unas de otras en una determinada propiedad. Este nivel de medición requiere que se establezca algún tipo de unidad de medida que pueda ser considerado por todos como una norma común y que sea repetible, esto es, que se pueda aplicar indefinidamente a los mismos objetos produciendo los mismos resultados. Para el caso, ello es lo que

22 Definición de Selltiz *et al.* (1965: 219).

ocurre en Física con las escalas de temperatura Celsius y Farenheit. En estas escalas, las unidades representan cantidades iguales de cambio en el volumen de una columna de mercurio bajo una determinada presión. Así, es posible afirmar por ejemplo que entre 0 C. y 10 C. existe la **misma diferencia** de temperatura que entre 10 C. y 20 C., vale decir, 10 grados Celsius en ambos casos. Pero subsiste todavía una limitación que deviene de la ausencia de un punto **cero absoluto** y que hace que las operaciones de multiplicación y de división carezcan de sentido, con lo que las proporciones resultan también carentes de significado. Tratando con una escala meramente intervalar como la de Celsius, no puede sostenerse que una temperatura de 20 C. sea el **doble** de una de 10 C. Debido a la arbitrariedad del punto cero en la escala Celsius —la temperatura de congelamiento del agua— no es posible formular este tipo de relaciones.[23] Para la mayoría de las variables en ciencias sociales no existen procedimientos de medición que garanticen tan siquiera la igualdad de intervalos; Thurstone no se equivocó al denominar a su escala "de intervalos aparentemente iguales". En este nivel se amplían sustancialmente los estadísticos aplicables, que incluyen ya a la media aritmética, la variancia, diversos modelos de regresión, y el coeficiente producto—momento de Pearson.

Nivel racional

También es llamado nivel 'de razones' o 'por cocientes': se alcanza al producir una escala de intervalos que cuenta con un punto cero absoluto. Es solamente en este nivel que se justifican afirmaciones del tipo «A pesa el doble que B», o bien «Yo gasté la mitad que vos». En Física, son numerosas las escalas definidas en este nivel: distancia en metros, intervalos de tiempo en segundos, medidas de peso, mas, velocidad, resistencia eléctrica, etc.

Empero, la distinción entre escalas de intervalos y de razones resulta irrelevante las más de las veces. Según Blalock, « esta distinción... es puramente académica, ya que es extremadamente difícil encontrar una escala legítima de intervalos que no sea al propio tiempo una escala de proporciones. Esto se debe al hecho de que, una vez establecida la magnitud de la unidad, es casi siempre posible concebir cero unidades, pese a que nunca podamos hallar un cuerpo que no posea largo o masa, u obtener una temperatura de cero absoluto. Así pues, prácticamente en todos los casos en que se dispone de una unidad, será legítimo emplear todas las operaciones de la aritmética, incluidas las raíces cuadradas, las potencias y los logaritmos» (1966: 27).

De acuerdo a esto, a todos los efectos prácticos podemos manejarnos reteniendo la distinción entre los tres niveles que aparecen en el cuadro sintético siguiente. Habrá que tener presente siempre que los niveles superiores presentan, además de las propiedades que les son propias, todas las

23 Esto se hace evidente mediante el sencillo expediente de expresar estas mismas temperaturas en grados Farenheit (mediante la fórmula: F= 1,8 C + 32) con lo que la relación no se mantiene: 68 F. no es ciertamente el "doble" de 50 F.

características que pertenecen a los niveles inferiores. Así, por ejemplo, a una variable de nivel intervalar como el ingreso mensual en pesos, se le puede aplicar —además de la media aritmética— la mediana y el modo aunque, por cierto, al precio de una gran pérdida de información.

Figura IV: Resumen de características de los niveles de medición

Nivel de medición	Operación básica	Medidas estadísticas admisibles	Ejemplos de variables
Nominal	Clasificación (a ≠ b)	Nº de casos, Modo, %, Q, Phi, Gamma, chi-cuadrada	Sexo, Estado civil, Afiliación política, Nacionalidad, Religión
Ordinal	Orden de las categorías (a > b)	Mediana, percentiles, cuartiles, Rho (Spearman)	NES, Nacionalismo, Religiosidad, Actitudes
Intervalar	Distancia entre categorías (a + b = c)	Media aritmética, Varianza, r (Pearson)	Edad en años, Ingreso en $, Tasa de suicidios

Algunos autores consideran que el subdesarrollo de las ciencias sociales obedece precisamente a la falta de medidas definidas en un nivel racional. Y es que, según Selltiz, «El descubrimiento de relaciones precisas entre características, tal como ha sido expresado en leyes numéricas, depende en buena parte de la existencia de escalas de razones» (1965: 225). Lo que se observa justamente en la historia de la ciencia es como se ha ido progresando de un nivel a otro. Como señala Stevens, «Cuando los hombres conocían la temperatura sólo mediante sensaciones, cuando las cosas eran sólo 'más calientes' o 'más frías' que otras, la temperatura pertenecía al tipo ordinal de escalas. Se convirtió en una escala de intervalo con el desarrollo de la termometría, y después que la Termodinámica utilizó la **razón** de expansión de gases para extrapolar a cero, se convirtió en una escala de razón».[24] Análogamente, la distinción entre los colores de los objetos que realizamos corrientemente en un nivel nominal,[25] desde la Física puede ser considerada como una escala de intervalos: se trata de la reflexión de la luz en diferentes longitudes de onda.

Ahora bien, cabe preguntarse si las ciencias sociales, hoy subdesarrolladas,

24 Cf. Artículo reproducido en Wainerman (Ed.), 1976: 64.
25 Aunque de modo variable en culturas diferentes.

deberían llegar con el tiempo —¿ cuánto más?— a mediciones intervalares de todas las propiedades fundamentales de sus objetos. ¿ O, por lo contrario, es que existe una característica intrínseca a dichos objetos que limita las posibilidades de cuantificación?[26] No existe una única respuesta a estos interrogantes. Si se atiende a la dimensión simbólica de lo social, a la importancia del significado atribuido por los actores a sus acciones, a la representación que se hacen de las relaciones sociales como un elemento que las constituye como tales, entonces puede carecer de sentido proponerse la medición de ciertas variables en un nivel más elevado que el que se plantean estos mismos actores. Así, para dar un ejemplo burdo, el valor peculiar que puede revestir el color 'rojo' en nuestra cultura política es algo que no mantiene ninguna relación con su longitud de onda.[27] Es decir, propiedades medibles en un nivel intervalar pueden funcionar socialmente en un nivel puramente nominal.[28]

Por otra parte, y para finalizar, también es frecuente que se produzca el caso inverso, cuando a propiedades nominalmente definidas se les puede atribuir un significado ordinal. Sin ir más lejos, es lo que ocurre con la variable 'ocupación', cuyas diferentes variedades distinguidas nominalmente son susceptibles de ser ordenadas según niveles de prestigio.[29] Análogamente, es corriente que en la confección de muchos índices se parta de propiedades clasificatorias a las que, a los efectos de la elaboración del índice, se les confiere un determinado valor numérico. Esto supondrá siempre ordenar las categorías de acuerdo con algún criterio, con lo que la propiedad dará lugar a la generación de una variable de nivel ordinal.

26 Según Passeron, tal subdesarrollo no responde a problemas de medición, sino a una conformación epistemológica insuperable de las disciplinas que denomina "ciencias del curso del mundo histórico" en las que «Las constataciones se dan en un contexto que puede ser *designado* pero no *agotado* por un análisis finito de las variables que lo constituyen y que permitirían razonar *ceteris paribus*.» (Passeron, 1991: 25).
27 Sólo pretendo ilustrar la escasa importancia de la "objetividad" de ciertas características, cuando su única función es de la ser soporte de significaciones para los actores. Una discusión en buen nivel de este problema más general puede encontrarse en la obra de Cicourel (1982: capítulo 1).
28 Integradas a sistemas clasificatorios que se estructuran en base a oposiciones simples: "rojo/blanco", "frío/caliente", etc.
29 Más estrictamente, deberíamos hablar en este caso de un único objeto —la 'ocupación'— del que podemos predicar varias propiedades (calificación, prestigio, etc.).